

研究报告

A Letter

鸡脂肪性状重要候选基因的连锁不平衡模式分析

李晓存* 卢国* 王守志 李辉**

东北农业大学动物科学技术学院 哈尔滨 150030

* 同等贡献作者

** 通讯作者 lihui@neau.edu.cn

摘要 对鸡脂肪性状候选基因进行连锁不平衡(linkage disequilibrium, LD)分析可以为理解鸡体脂沉积的遗传学机制提供重要信息,并且可以提高脂肪性状相关分子遗传标记的检出效率。本研究以东北农业大学肉鸡高、低脂双向选择品系第 8 世代肉仔鸡为实验材料,通过直接测序、高效液相色谱(DHPLC)、单链构象多态性(PCR-SSCP)、限制性片段长度多态性(PCR-RFLP)、长度多态性(PCR-LP)等方法对过氧化物酶增殖体激活受体(*PPAR-γ*)、载脂蛋白 B(*ApoB*)、解耦联蛋白(*UCP*)以及其他 10 个鸡脂肪性状重要候选基因进行多态性检测,计算候选基因标记间 LD 参数 $|D'|$ 与 r^2 ,分析这些候选基因标记间的 LD 模式并探讨其规律。初步发现了以下特征:1、在各候选基因内,标记间的 $|D'|$ 与 r^2 值并不匹配,取值区间差异很大, $|D'|$ 值普遍大于 r^2 ;另外,在有些基因内的相同标记间 $|D'|$ 、 r^2 值相差很大,甚至在 $|D'|=1$ 时 r^2 接近于 0。2、在不同基因甚至在同一基因内不同区域 LD 水平差异很大且物理距离与 LD 关系并不成线性关系。3、在有些基因内,如 *ApoB* 基因,相间位点间的 $|D'|$ 和 r^2 值大于相邻位点间的值。4、有些基因,如 *PPAR-γ*, 相同标记间 LD 水平在高、低脂系间存在很大差异。本研究结果提示,在群体遗传学中进行候选基因 LD 分析时,LD 参数 r^2 的应用价值比 $|D'|$ 更大。

关键词 鸡 脂肪性状 候选基因 连锁不平衡模式 $|D'|$ r^2

Linkage Disequilibrium Patterns of Polymorphisms in Candidate Genes for Fatness in Chickens

Li Xiaocun* Hu Guo* Wang Shouzhi Li Hui**

College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

* These authors contributed equally to this work

** Corresponding author, lihui@neau.edu.cn

DOI: 10.3969/j.issn.1674-7968.2011.02.012

Abstract Investigating patterns of linkage disequilibrium(LD) is helpful for a more in-depth understanding of the genome-wide mechanisms for the genetic variation response to the natural and artificial selection in model and domestic animals. LD analysis for the candidate genes on fatness in chickens can provide more important information about the genetic control of fat deposition and make the molecular markers associated with fatness more powerful. In this study, polymorphisms of peroxisome proliferator-activated receptor- γ (*PPARγ*), apolipoprotein B (*ApoB*), uncoupling protein (*UCP*) and other 10 candidate genes on chicken fatness among individuals were detected by DNA sequencing, denaturing high performance liquid chromatography(DHPLC), single-strand conformation polymorphism(SSCP), restriction fragment length polymorphism(RFLP) and labelled primers(LP) methods. LD investigation among the polymorphisms of the candidate genes was performed in the 8th

DOI: 10.3969/j.issn.1674-7968.2011.02.012

基金项目:本研究由国家高技术研究发展计划(863)项目(No.2006AA10A120)、国家重点基础研究发展规划(973)项目(No.2006CB102105)和现代农业产业技术体系建设专项(No.nycyt-42)共同资助

收稿日期 2010-04-13 接受日期 2010-07-09

generation populations of the Northeast Agricultural University broiler lines divergently selected for abdominal fat content, two common measures of LD, $|D'|$ and r^2 , were used in this study. The results showed that the value of $|D'|$ and r^2 often did not match very much within the candidate genes. The value of $|D'|$ usually was larger than value of r^2 , sometimes even $|D'|=1$, while r^2 was close to zero. Furthermore, in different genes, even within the same gene, LD patterns of different regions were usually different, and there was no a linear correlation between physical distance and the value of $|D'|$ or r^2 . In addition, in some genes, such as *ApoB*, the values of $|D'|$ and r^2 between nonadjacent polymorphisms were much larger than those between the adjacent polymorphisms. Finally, there were different LD patterns of polymorphisms in some genes, such as *PPAR-γ* between lean and fat chicken lines. The results in this study also suggested, as a LD measure parameter, r^2 is more useful than $|D'|$ in population genetics analysis.

Keywords Chicken, Fatness, Candidate gene, Linkage disequilibrium pattern, $|D'|$, r^2

脂肪的过度积累是引起机体肥胖的主要原因之一(王建平等, 2009)。脂肪性状是一类重要的经济性状, 同时也是典型的复杂性状。人类、哺乳动物和禽类脂肪代谢都涉及众多的信号转导通路和代谢途径, 其中包括很多重要的转录调控因子和功能基因, 它们在脂肪合成、运输、贮藏、分解等重要的代谢过程中发挥重要的功能。

连锁不平衡亦称为配子相不平衡(gametic phase disequilibrium)、配子不平衡(gametic disequilibrium)或等位基因关联(allelic association) 指的是一个群体内不同座位等位基因之间的非随机关联, 包括两个标记间或两个基因/数量基因座(quantitative trait loci, QTL)间或一个基因/QTL与一个标记座位间的非随机关联(王荣焕等, 2007)。目前, 连锁不平衡(linkage disequilibrium, LD)分析在畜禽群体基因精细定位等研究中得到了广泛应用。Meuwissen等(2002)用连锁分析与连锁不平衡分析相结合的方法将影响牛双胞胎比率的QTL精细定位到5号染色体CSSM22~IL STS66两个标记之间。Sutter等(2004)通过对狗基因组的连锁不平衡分析发现狗的连锁不平衡程度比人的要大的多, 在基因定位研究中, 达到相同的定位精度所需要的标记数量较之人要少的多。对日本肉牛的LD研究发现, 利用LD对重要经济性状的QTL进行精细定位是一个非常有效的工具(Odani et al., 2006)。Abasht等(2009)在商业蛋鸡品系中通过连锁不平衡对影响蛋品种性状的标记测定发现, 如果与QTL相关的标记处在一个小的物理距离范围内, 且两世代都处于强的连锁不平衡状态, 那么这些与QTL紧密连锁的标记可以应用于标记辅助选择改良性状。尽管应用LD进行重要经济性状QTL精细定位研究取得了很大的进展, 然而基于候选基因内LD模式研究目前还鲜有报道。对鸡脂肪

性状候选基因进行LD模式分析不仅有助于寻找功能性变异位点从而提高脂肪性状相关分子遗传标记选择的效率, 而且可以为最终理解鸡体脂沉积的遗传学机制提供重要信息。本研究通过对鸡脂肪性状重要候选基因内标记间LD模式及规律研究, 以期揭示重要候选基因多态性位点之间内在联系, 为低脂肪肉鸡选育提供依据。

1 结果与分析

1.1 连锁不平衡度量值($|D'|$ 、 r^2)分析

根据连锁不平衡分析结果(表1)发现, 在各候选基因内, 标记间的 $|D'|$ 与 r^2 值并不匹配, 取值区间差异很大。就具体数值来讲, $|D'|$ 值普遍大于 r^2 ; 在有些基因内相同标记间, 甚至在 $|D'|=1$ 时, r^2 接近于0。例如在高脂系内*LFABP*基因的g.-204G>A与g.3628A>G位点间LD度量值为 $|D'|=1$ 而 $r^2=0.009$ 。在有些基因内相间位点间的 $|D'|$ 和 r^2 值大于相邻位点间的值, 例如, 在高脂和低脂两个品系内*ApoB*基因g.-112A>G与g.34202indel(9 bp)位点间的 $|D'|$ 、 r^2 值均大于g.-112A>G与g.21751T>G间的相应数值。

1.2 连锁不平衡程度分析

根据连锁不平衡分析结果(表1)发现, 在不同基因甚至在同一基因内不同区域LD水平差异很大, 且物理距离与LD关系并不成比例; 有些基因, 相同标记间LD水平在高、低脂系间存在很大差异, 就具体取值来说, 高脂系LD取值高于低脂系, 如*PPAR-γ*基因在高脂系中LD水平明显高于低脂系。

2 讨论

2.1 连锁不平衡度量值($|D'|$ 、 r^2)分析

自Lewontin和Kojima(1960)提出LD的概念以

表 1 各候选基因内连锁不平衡程度分析结果
Table 1 Linkage disequilibrium analysis of candidate genes

基因 Gene	SNP 组合 SNP Combination	物理距离 /bp Distance	品系 Breed	r^2	$ D' $	品系 Breed	r^2	$ D' $
<i>AFABP</i>	No.1-No.2	2162	L	--*	--	H	0.080	1.000
	No.1-No.3	4148		--	--		0.001	1.000
	No.2-No.3	1986	L	0.004	0.145	H	0.001	1.000
<i>HFABP</i>	No.4-No.5	4272		0.029	0.381		0.023	0.182
	No.4-No.6	4931		0.183	0.511		0.044	0.289
	No.4-No.7	5627		0.160	0.883		0.025	0.451
	No.5-No.6	659		0.193	0.833		0.012	0.354
	No.5-No.7	1355		0.026	0.805		0.007	1.000
<i>IFABP</i>	No.6-No.7	696		0.110	0.881		0.006	0.312
	No.8-No.9	1166	L	0.001	0.075	H	0.002	0.117
	No.8-No.10	1580		0.074	0.900		0.067	0.785
	No.8-No.11	3829		0.217	0.862		0.124	0.715
	No.9-No.10	414		0.006	0.717		0.015	1.000
<i>LFABP</i>	No.9-No.11	2663		0.035	1.000		0.008	0.474
	No.10-No.11	2249		0.283	0.963		0.436	0.97
<i>UCP</i>	No.12-No.13	3832	L	0.240	0.971	H	0.009	1.000
<i>ApoB</i>	No.14-No.15	76	L	0.037	0.289	H	0.280	0.793
	No.14-No.16	1354		0.305	0.750		0.398	0.668
	No.15-No.16	1278		0.229	0.969		0.413	1.000
<i>ACCa</i>	No.17-No.18	21863	L	0.444	0.953	H	0.573	0.931
	No.17-No.19	34314		0.702	0.953		0.908	1.000
	No.18-No.19	12451		0.605	0.974		0.546	0.868
<i>IGFI</i>	No.20-No.21	61876	L	0.132	0.531	H	0.343	0.669
	No.22-No.23	39076	L	0.929	0.964	H	0.748	1.000
	No.22-No.24	48388		0.947	0.982		0.497	1.000
<i>MC3R</i>	No.23-No.24	9312		0.877	0.945		0.665	1.000
	No.25-No.26	276	L	0.006	0.303	H	0.073	0.673
<i>MC4R</i>	No.27-No.28	307	L	0.430	0.899	H	0.227	0.665
<i>PPAR-γ</i>	No.29-No.30	527	L	0.595	0.797	H	0.889	0.949
	No.29-No.31	1963		0.187	1.000		0.512	1.000
	No.29-No.32	3671		0.177	0.883		0.575	0.979
	No.30-No.31	1166		0.062	0.553		0.464	0.957
	No.30-No.32	3144		0.056	0.482		0.523	0.940
	No.31-No.32	1978		0.627	0.879		0.713	0.924
<i>IGFBP2</i>	No.33-No.34	2931	L	0.587	0.952	H	0.316	0.871
<i>MDI</i>	No.35-No.36	69	L	0.025	0.808	H	0.131	0.525

* 因位点无多态而造成的缺失值 ; *AFABP* : 脂肪型脂肪酸结合蛋白基因 ; *HFABP* : 心脏型脂肪酸结合蛋白基因 ; *IFABP* : 小肠型脂肪酸结合蛋白基因 ; *LFABP* : 肝脏型脂肪酸结合蛋白基因 ; *UCP* : 解耦蛋白基因 ; *ApoB* : 载脂蛋白 B 基因 ; *ACCa* : 乙酰辅酶 A 羧化酶 α 基因 ; *IGFI* : 类胰岛素生长因子 1 基因 ; *MC3R* : 黑素皮质素受体 3 基因 ; *MC4R* : 黑素皮质素受体 4 基因 ; *PPAR-γ* : 过氧化物酶增殖体激活受体 γ 基因 ; *IGFBP2* : 胰岛素样生长因子结合蛋白基因 2 ; *MDI* : 苹果酸脱氢酶 1 基因 L : 低脂系 H : 高脂系

*: The sites without data are no polymorphism; *AFABP*: Gene of adipocyte fatty acid binding protein; *HFABP*: Gene of heart fatty acid binding protein; *IFABP*: Gene of intestinal fatty acid binding protein; *LFABP*: Gene of liver fatty acid-binding protein; *UCP*: Gene of uncoupling protein; *ApoB*: Gene of apolipoprotein B; *ACCa*: Gene of acetyl-CoA carboxylase α; *IGFI*: Gene of insulin-like growth factor 1; *MC3R*: Gene of melanocortin 3 receptor; *MC4R*: Gene of melanocortin 4 receptor; *PPAR-γ*: Gene of peroxisome proliferator-activated receptor-γ; *IGFBP2*: Gene of insulin-like growth factor-binding protein 2; *MDI*: Gene of malate dehydrogenase; L: Lean line; H: Fat line

来,经过多年的发展,研究者提出了多种 LD 的统计量(Devlin and Risch, 1995)。一般说来,各种 LD 统计量的目的都是比较实际观测到的单倍型频率与配子随机分离时单倍型期望频率之间的差异,进而估算位点间 LD 的水平。在具体分析中,LD 统计量的应用需要根据位点的数目与标记的性质来确定,对于双等位基因的座位,如 SNPs 或插入/缺失突变, $|D'|$ 和 r^2 是应用最为普遍的统计量(Jorde, 2000)。为了通过 LD 统计量(如 $|D'|$ 、 r^2)的数值界定各位点间的 LD 水平,就必须理解这些统计量的遗传学意义。一般说来, $|D'|$ 和 r^2 反映了 LD 的不同方面, $|D'|$ 仅包括了重组史,而 r^2 在包括了重组史同时也包括了突变史以及其诸如自然或人工选择、基因转换、遗传漂变、群体扩张、群体混入等可以影响 LD 水平的历史事件(Slatkin, 2008)。已有研究表明,在长期的进化历程中,重组是影响 LD 的主要因素(Shifman et al., 2003),但在短期的育种过程中选择引起的等位基因频率的定向改变以及遗传漂变等则为影响 LD 水平的主要因素(Slatkin, 2008)。

本研究中发现在各候选基因内,各标记间的 $|D'|$ 值普遍大于 r^2 ; 在有些基因相同标记间, $|D'|$ 、 r^2 值相差很大,甚至在 $|D'|=1$ 时 r^2 接近于 0。这是由于 $|D'|$ 、 r^2 不同的遗传学意义造成的。 $|D'|$ 度量值在小样本中存在着夸大效应(Ardlie et al., 2002)。而 $|D'|=1$ 时 r^2 接近于 0 的现象,在人类复杂疾病致病基因 LD 分析中也有类似的发现,引起该现象的一个可能原因是 r^2 比 $|D'|$ 对等位基因频率更为敏感(Devlin and Risch, 1995; Nothnagel et al., 2002)。本研究中在高脂系内 *LFABP* 基因的 g.-204G>A 与 g.3628A>G 位点间 LD 度量值 $|D'|=1$, 而 $r^2=0.009$ 。g.3628A>G 位点最小等位基因频率就很低,仅为 5.8%。其他出现 $|D'|=1$ 时 r^2 接近于 0 现象的标记也存在类似现象。此外,东北农业大学肉鸡高、低脂双向选择品系两系都是由相同基础群闭锁选育得到的,选育时期相较于进化来说非常短,重组史差异不大,引起 LD 的主要原因为选择和遗传漂变,这是导致 $|D'|$ 与 r^2 值不匹配的重要原因之一。

本研究还发现,在有些基因内相间位点间的 r^2 和 $|D'|$ 值大于相邻位点间的相应值。需要明确指出的是,这种现象并不能简单理解为相间位点的 LD 水平高于相邻位点,产生该情况的原因主要是由 r^2 、 $|D'|$ 的取值特性不同造成的(VanLiere and Rosenberg, 2008)。实际上, $|D'|$ 与 LD 强度之间并不成线性的定量关系,而是呈现一种分区段的定性关系,不同研究

中的 LD 水平不能通过直接比较 r^2 、 $|D'|$ 值来评价(Wall and Pritchard, 2003)。当在检测位点间观察不到任何重组事件的时候 $|D'|=1$, 此时称为完全连锁不平衡(complete LD),能观察到 2 或 3 种单倍型;当观测到 3 种单倍型时, $|D'|=1$, $r^2 \neq 1$, 这表示两个位点不完全连锁,但没有重组的证据。这种 LD 的形成反映出两个位点有相同的重组历史,但突变及其他影响 LD 水平的历史不同,这暗示着在特定群体中这两个多态性位点来源的系谱可能存在差异(Flint-Garcia et al., 2003)。 $|D'|=0$ 说明两个位点完全独立, $|D'|<1$ 说明两个位点间发生了历史重组,完全 LD 被破坏,但是,当 $|D'|<1$ 时的一些中间值的意义是很难解释的。只有在统计学上显著接近 1 的 D' 值,一般认为 $|D'|>0.8$ (也有较宽的标准认为 $|D'|>0.7$ (Tiret et al., 2002)),才能提示这两个位点之间有较小的历史重组,而中等程度的 $|D'|$ 值不能用来比较不同研究中的 LD 程度(Pritchard and Przeworski, 2001),甚至在模拟情况下,相同距离的成对位点间的 D' 值也是高度可变的(Hudson, 1985; Hudson, 2001)。

2.2 连锁不平衡程度分析

对人类复杂疾病易感基因的研究结果表明,不同基因甚至在同一基因内不同区域标记间 LD 水平差异很大且物理距离与 LD 关系并不成比例(Long, 2004)。本研究也获得了类似的结果,能够导致这类现象产生的素因有很多,以下为最主要的两种原因:第一,鸡基因组不同染色体以及染色体的不同区段的重组率和 LD 模式都有很大差异(Aerts et al., 2007; Andreescu et al., 2007);第二,LD 水平与位点所处基因不同区域也有密切关系。Eberle 等(2006)的研究表明,基因区域标记间的 LD 水平显著高于非基因区域,LD 水平在染色体基因簇的侧旁开始降低直至连锁平衡。而本研究中,有的标记间的物理距离相对较近,但是二者已处于基因的侧翼;而有些距离较远的位点因其都处于基因的编码区,受自然选择压力的影响,后者的 LD 水平很大程度上倾向于高于前者。

此外,对人类基因组及重要功能基因的 LD 研究表明,自然选择会对重要功能基因位点周围的 LD 模式产生影响,受到正向进化选择的位点附近的 LD 的水平显著加强、宽度拓展延伸(Sabeti et al., 2002)。对家养动物的人工选择驯化与自然选择过程相近似,选择强度更大,取得遗传进展也更为明显。对鸡基因组的 LD 研究也表明不同品种、品系鸡染色体都有其相对独特的 LD 模式,可以作为区分这些品

种、品系遗传特征的依据 (Aerts et al., 2007; Andreescu et al., 2007; Rao et al., 2008)。这提示品种、品系形成过程中, 遗传选择可以显著地改变染色体的 LD 模式, 是影响群体 LD 模式的重要因素之一。本研究中肉鸡高脂系与低脂系群体都是从相同的基础群中, 经闭锁双向选育获得的, 且饲养管理条件相同。本研究中发现有些基因相同位点间 LD 水平在高脂系与低脂系间的差异很大, 如 *PPAR-γ* 基因在高脂系中 LD 水平明显高于低脂系, 表明 *PPAR-γ* 基因对这种双向选择产生了不同的应答方式。这种结果提示我们, 在小的育种群体内, 遗传选择可能是改变候选基因的 LD 模式的重要原因之一。

2.3 连锁不平衡遗传学分析

根据对本研究结果的分析还发现, 在群体遗传学分析中 r^2 的应用价值比 $|D'|$ 更大, 这是由 r^2 的遗传学意义和数理特性决定的 (Devlin and Risch, 1995; Flint-Garcia, 2003; Slatkin, 2008)。当 r^2 等于 0 时也说明两个位点之间是完全独立的; 不过 $r^2=1$ 与 $|D'|=1$ 的意义则有所不同, $r^2=1$ 称为完美连锁 (perfect LD), 此时两个位点经历了相同的突变、重组以及其他影响 LD 水平的历史事件, 两个位点中的一个位点某个等位基因的出现完全预示着另一个位点相应等位基因的出现, 这时候两个位点组成的 4 种可能单倍型中仅表现 2 种单倍型, 即 $r^2=1$, 同时必然有 $|D'|=1$ 。另外, 与 $|D'|$ 比较 r^2 具有更充分的群体遗传学理论基础和统计学优点。第一, r^2 的期望值和有效群体大小和重组系数相关 $E(r^2) = 1/(1+4N_eC)$, 其中 N_e 是有效群体的大小 (effective population size) (Pritchard and Przeworski, 2001), C 是重组系数; 第二, r^2 有很好的取样特性 (Przeworski and Wall, 2001), 在 case-control 型关联分析时, 样本量和 r^2 的乘积就是所观察到的关联水平尾概率对应的卡方值。在检测 SNPs 和致病位点之间的关联时, 如果要达到同样的统计效力, 所需用的样本量要增大 $1/r^2$ 倍, 依据这个特性可以利用 r^2 统计量可以进行 tagSNP 的筛选 (Kruglyak, 1999; Pritchard and Przeworski, 2001; Wall and Pritchard, 2003)。第三, 在全基因关联分析中, 标记间的 r^2 值与 QTL 的捕捉效率有密切关系。有研究表明, 当相邻标记间 $r^2>0.2$ 时, QTL 的捕捉效率能够达到 0.8 (Calus and Veerkamp, 2007)。第四, 与 $|D'|$ 相比, 在同样长度的染色体范围内, r^2 往往要更低。相应地, 当 $r^2>0.1$ (Kruglyak, 1999; Nakajima et al., 2002) (也有较严格的标准认为 $r^2>0.33$ (Ardlie et al., 2002;

Shifman et al., 2003)), 即认为两个位点间处于强 LD 状态, r^2 的这种取值特性能够帮助我们获得更准确的单倍型块, 进而提高 QTL 定位的精度 (Pritchard and Przeworski, 2001)。总之, 在群体遗传学分析中 r^2 比 $|D'|$ 有更为重要的应用价值。但是, 不可忽视的一点是, 由于 r^2 并不能和重组率直接联系起来, 因此, 在有些更关注重组率等历史事件的应用领域如进化遗传学等, r^2 的应用受到限制, 一般只作为 $|D'|$ 的辅助 (Devlin and Risch, 1995; Flint-Garcia, 2003)。

综上所述, 本研究通过对 13 个重要的鸡脂肪性状候选基因 LD 模式及规律研究发现, 在一般意义上, LD 水平与候选基因所处染色体位置, 标记所处基因的位置差异等因素有重要关系; 基于本研究所用群体特点, 我们认为遗传选择可能是影响该群体与脂肪性状有关的候选基因内 LD 模式的重要因素之一。本研究有助于更为深入地理解候选基因的 LD 模式及其遗传规律, 为揭示重要候选基因功能位点之间内在联系, 理解复杂性状遗传机理提供相应的参考依据, 进而为低脂肉鸡的选育提供参考。

3 料与方法

3.1 实验鸡群

以东北农业大学选育的肉鸡高、低脂双向选择品系第八世代仔鸡 398 只为实验材料, 其中高脂系为 171 只, 低脂系为 227 只。实验鸡群按常规方法进行饲养管理。47 日龄时翅静脉采血, EDTA 抗凝, 酚-氯仿抽提 DNA 之后, TE 溶解, -20°C 保存。

3.2 候选基因选择与多态性检测

本课题组已对脂肪酸结合蛋白家族 (*A-FABP*、*H-FABP*、*I-FABP*、*L-FABP*)、解耦蛋白 (*UCP*)、载脂蛋白 B (*ApoB*)、乙酰辅酶 A 羧化酶 a (*ACCα*)、类胰岛素生长因子 (*IGF1*)、黑素皮质素受体 3 (*MC3R*)、黑素皮质素受体 4 (*MC4R*)、过氧化物酶增殖体激活受体 γ (*PPARγ*)、胰岛素样生长因子结合蛋白 (*IGFBP2*) 和苹果酸脱氢酶 (*MD1*) 等 13 个基因作为影响鸡脂肪代谢的候选基因进行了多态性检测。分布于这 13 个基因的共 36 个多态位点用于 LD 研究。根据其在基因上位置从 5' 到 3' 按顺序编号, 并且根据 GenBank 相关信息 (<http://www.ncbi.nlm.nih.gov/Genbank>) 与鸡基因组测序结果 (www.genome.ucsc.edu) 分别对以上位点设计引物在东北农业大学肉鸡高、低脂双向选择品系第 8 世代个体进行 PCR 扩增 (10 μL 体系) 50

ng/ μ L 基因组 DNA 1 μ L、10 \times PCR buffer 0.5 μ L、10 mmol/L dNTP 0.4 μ L、10 μ mol/L 上下游引物各 0.25 μ L、5 U/ μ L *Taq* DNA 聚合酶 0.1 μ L、ddH₂O 7.5 μ L；PCR 扩增条件为 94 $^{\circ}$ C 7 min, 94 $^{\circ}$ C 30 s, 51~64 $^{\circ}$ C 30 s 28~33 个循环 (不同引物退火温度和循环数不同), 72 $^{\circ}$ C 30 s, 72 $^{\circ}$ C 延伸 7 min。PCR 扩增产物通过 DHPLC(仪器自动吸取 PCR 产物 5 μ L, 以变性温度将样本在 DNA 分离柱中洗脱分型)、PCR-SSCP(1 μ L 的 PCR 产物、4 μ L 的变性上样缓冲液, 利用 14% 的聚丙烯酰胺凝胶电泳检测观察结果)、PCR-RFLP(5 μ L 的 PCR 产物、3 U 内切酶、2 μ L buffer 和 12.9 μ L 的 ddH₂O 混合、37 $^{\circ}$ C 过夜 利用琼脂糖凝胶电泳检测观察结果)、PCR-LP (1 μ L 的 PCR 产物、3 μ L 的非变性上样缓冲液 利用 14% 的聚丙烯酰胺凝胶电泳检测观察结果)等方法进行多态性检测。各候选基因名称、定位以及多态性位点位置、PCR 扩增引物和退火温度、分型方法及其最小等位基因频率等相关信息见表 2。

3.3 统计分析

本研究采用 LD 分析中较为常用的统计量 $|D'|$ 、 r^2 进行 LD 水平估计(Devlin and Risch, 1995)。假设两个位点 共形成 4 种单倍型 单倍型的频率分别表示为 f_{11} 、 f_{12} 、 f_{21} 、 f_{22} 。而 $f_{1+} = f_{11} + f_{12}$ 、 $f_{+1} = f_{11} + f_{21}$ 、 $f_{2+} = f_{21} + f_{22}$ 、 $f_{+2} = f_{12} + f_{22}$ 。LD 的基本统计量为 $D = f_{11}f_{22} - f_{12}f_{21}$ 。

则有：

$$\text{当 } D > 0 \text{ 时 } D' = \frac{D}{\min(f_{1+}f_{+2}, f_{1+}f_{2+})} ;$$

$$\text{当 } D < 0 \text{ 时 } D' = \frac{D}{\min(f_{1+}f_{+1}, f_{+2}f_{2+})} ;$$

$$r^2 = \frac{D^2}{f_{1+}f_{2+} \cdot f_{+1}f_{+2}} ;$$

利用第 8 世代高、低脂两系群体个体基因型信息分别进行候选基因标记间的 LD 分析, 计算同一基因多态性位点两两之间的 $|D'|$ 、 r^2 值, 统计软件为 SAS9.1.3 中的 allele 过程(SAS Inst. Inc., Cary, NC)。

表 2 各候选基因及突变位点基本信息

Table 2 The basic information of candidate genes and mutant sites

染色体	基因	位点编号	突变位置	引物信息	退火温度/ $^{\circ}$ C	分型	品系	最小等位基因频率
Chromosome	Gene	No. of SNPs	Domain	Primer	T_m	Typing	Breed	MAF
chr2	<i>AFABP</i>	No.1	5'UTR	5'-GCCAAGGAAGTAGGATATA-3'	55.6	LP ¹	L	0.000
			g.-433indel(7bp)	5'-TTTCAGTCCTCCATATTGTAT-3'			H	0.272
		No.2	exon3	5'-GGTCCAAAGCACCCCTGATGA-3'	56.4	DHPL	L	0.230
			g.1729G>A	5'-CCATCCACCACTTTCCTCTT-3'			H	0.311
		No.3	3'UTR	5'-CGAAAGAGCATGAGGAAG-3'	59.1	SSCP ³	L	0.055
			g.3265T>C	5'-GCCACAACAAATTAGGC-3'			H	0.007
chr23	<i>HFABP</i>	No.4	5'UTR	5'-CAGTGGCATAAGGGTCAAAG-3'	55.6	<i>Dra</i>	L	0.327
			g.-2438T>C	5'-CCCAGTGCACATATGTTTC-3'			H	0.357
		No.5	intron2	5'-AGTGGGTTGCAGGGACTCACA-3'	55	LP	L	0.292
			g.1834indel(14bp)	5'-GGGCTCTGCATGACTTAAGA-3'			H	0.291
		No.6	intron2	5'-TGAGCCCCTGTCACGCTAAC-3'	60.1	<i>Nco</i>	L	0.269
			g.2493C>T	5'-AGTTCAGTGCTGGAGCTGGC-3'			H	0.493
Chr23	<i>HFA</i>	No.7	3'UTR	5'-TTGGGTTGAGGGCGTTAT-3'	64	LP	L	0.087
			g.3189indel(8bp)	5'-CAGATAAAGGGGACGGTGAT-3'			H	0.018
chr4	<i>IFABP</i>	No.8	5'UTR	5'-ACTTAGCAGCACATCAAC-3'	57	<i>Tap</i>	L	0.462
			g.-561A>C	5'-CAGTCACTAACCCATTTCAT-3'			H	0.454
		No.9	intron2	5'-GAGGGAAATTGCTGGTC-3'	55	<i>Pvu</i>	L	0.032
			g.605C>T	5'-TTTGCATTTTATTGTGACTTTTT-3CTGCAG-3'			H	0.036
		No.10	intron2	5'-GATGGCATTAAACGGTACTTGGA-3'	55	<i>SnaB</i>	L	0.072
No.11	3'UTR	5'-TCTTGGCTTCAACTCCTTCGTAC-3'			H	0.085		
		5'-CTGTCAAGAAGAGCACCAG-3'	64	<i>Mae</i>	L	0.496		
chr4	<i>LFABP</i>	No.12	5'UTR	5'-CATTGTGTTATGGATGCCACG-3'	55	SSCP	L	0.425
				5'-TAAGGACCTTTTGCCCTAA-3'			H	0.122
		No.13	3'UTR	5'-CTAGGGACCTTAAACAGG-3'	61	<i>Asu</i> C2	L	0.253
				g.3628A>G	5'-CACTTTGGCTAAGTCTTCG-3'			H

chr1	UCP	No.14	intron2	5'-GCCCCAAGTGCTCAACCCCT-3'	60.2	Eco72	L	0.243
			g.1240C>A	5'-CTGGTGCTATGAGTCTGTTCC-3'				
		No.15	exon3	5'-AGGGCTGCTGGCGCGGCTGCTT-3'	59.8	Hha	L	0.419
			g.1316C>T	5'-CTGGTGCTATGAGTCTGTTCC-3'				
No.16	3'UTR	5'-GCTCCTGGAACGTGGTGA-3'	56	Hae	L	0.413		
	g.2594C>A	5'-GCTGCCTTTGGTCCCTCT-3'						
chr3	ApoB	No.17	5'UTR	5'-AACCCAACAACCTGGGATAG-3'	51	Bsp1407	L	0.431
			g.-112A>G	5'-TGACTTGGACCTTCATTTGTA-3'				
		No.18	exon26	5'-CATATTTCTAATGGCATCCAG-3'	52	Acy	L	0.268
			g.21751T>G	5'-TTCCCAGCGTTATTTCCG-3'				
No.19	3'UTR	5'-AAGGTCATATTTGTAAGCC-3'	52	LP	L	0.365		
	g.34202indel(9bp)	5'-TTCTTTTCTGCCATTTGATAC-3'						
chr19	ACCa	No.20	exon19	5'-TAGGTGGATGGTTGGGCTTGA-3'	56	Mwo	L	0.346
			g.21726G>A	5'-CCCCATCCTCCACCACATG-3'				
		No.21	exon45	5'-TTTGGTGCCATACATCGTG-3'	54.4	Taq	L	0.490
chr1	IGF1	No.22	5'UTR	5'-CACAGCCACCCGAAAGT-3'	57	Hinf	L	0.130
			g.-366A>C	5'-AGAAATCACAAAAGCAGCAC-3'				
		No.23	exon3	5'-TACACATCTACCACTGTCAT-3'	55	SSCP	L	0.130
			g.38710G>A	5'-TCCTCAGGTCACAACCTCT-3'				
No.24	3'UTR	5'-TTGGCTAATGGCTACTTTG-3'	55	Bsp119	L	0.150		
	g.48022C>T	5'-GTGCGTCATATTATCTTTTCG-3'						
chr20	MC3R	No.25	exon1	5'-TGGCTGTGGCAGATATGTTA-3'	56	SSCP	L	0.056
			g.273C>T	5'-AGTTAGGTACCTGTCAATGGCT-3'				
		No.26	exon1	5'-GACGACCACTTCATTCAG-3'	57.6	Dde	L	0.496
chr2	MC4R	No.27	exon1	5'-GAATTCACCCAGCATCG-3'	55	SSCP	L	0.208
			g.8A>G	5'-GAGGTTCTTGTCTTGGCTAT-3'				
		No.28	exon1	5'-GCAATAGCCAAGAACAAGA-3'	55	SSCP	L	0.327
			g.315G>T	5'-GCTCTGTGCGTCTGTATCTG-3'				
chr12	PPAR-γ	No.29	5'UTR	5'-TCTGGAAGGAGTGGAGTC-3'	55.3	LP	L	0.378
			g.-1784_-1768	5'-ATGCCCGATAAAGTAGAG-3'				
		No.30	5'UTR	5'-GGGCGGTCACTGTTGTAC-3'	52.6	SSCP	L	0.391
			g.-1241G>A	5'-CATTTGGATGACAGAACTTAT-3'				
		No.31	5'UTR	5'-GATGAAGGGAGACGAGAC-3'	60	Stu	L	0.235
No.32	exon2	5'-GTGCAATCAAAATGG AGCC-3'	55	SSCP	L	0.403		
chr7	IGFBP2	No.33	intron2	5'-CCAGACAAAATCAACCTTAGC-3'	58	Eco72	L	0.147
			g.2245C>T	5'-AATGGTGCTTTGCTGCTTA-3'				
		No.34	3'UTR	5'-GGGACTGCTTTCCAATAG-3'	54.0	SSCP	L	0.212
chr7	MD1	No.35	5'UTR	5'-TCCTCCAGTTCAATACAAGC-3'	62	Sph	L	0.081
			g.-4245C>T	5'-ATCAGTTCTGTCTGTGCC-3'				
		No.36	5'UTR	5'-ATTCCTTGACCGCTATGC-3'	58	SSCP	L	0.445
			g.-4176C>T	5'-TGCCCAGGAAGAAGAGCC-3'				

基因各缩写注释同表 1 ;LP :长度多态 ;DHPLC :高效液相色谱 ;SSCP :单核苷酸构象多态 ;其余位点分型方法均为限制性片段长度多态性(RFLP)分型所用内切酶名称 ;L :低脂系 ;H :高脂系

Notes of genes see Table 1; LP is length polymorphisms; DHPLC is denaturing high performance liquid chromatography; SSCP is single-strand conformation polymorphism; The others are incision enzymes used in restriction fragment length polymorphism (RFLP); L :Lean line; H :Fat line

参考文献

- Abasht B., Sandford E., Arango J., Settar P., Fulton J.E., O'Sullivan N.P., Hassen A., Habier D., Fernando R.L., Dekkers J.C., and Lamont S.J., 2009, Extent and consistency of linkage disequilibrium and identification of DNA markers for production and egg quality traits in commercial layer chicken populations, *BMC Genomics*, 10:1186/1471~2164
- Aerts J., Megens H.J., Veenendaal T., Ovcharenko I., Crooijmans R., Gordon L., Stubbs L., and Groenen M., 2007, Extent of linkage disequilibrium in chicken, *Cytogenetic and Genome Research*, 117:338~345
- Andrescu C., Avendano S., Brown S.R., Hassen A., Lamont S.J., and Dekkers J.C.M., 2007, Linkage disequilibrium in related breeding lines of chickens, *Genetics*, 177:2161~2169
- Ardlie K.G., Kruglyak L., and Seielstad M., 2002, Patterns of linkage disequilibrium in the human genome, *Nature Reviews Genetics*, 3:299~309
- Calus M.P.L., and Veerkamp R.F., 2007, Accuracy of breeding values when using and ignoring the polygenic effect in genomic breeding value estimation with a marker density of one SNP per cM, *Journal of Animal Breeding and Genetics*, 124:362~368
- Devlin B., and Risch N., 1995, A comparison of linkage disequilibrium measures for fine-scale mapping, *Genomics*, 29(2):311~322
- Eberle M.A., Rieder M.J., Kruglyak L., and Nickerson D.A., 2006, Allele frequency matching between SNPs reveals an excess of linkage disequilibrium in genic regions of the human genome, *PLoS Genetics*, 2:e142
- Flint-Garcia S.A., Thornsberry J.M., and Buckler E.S. 4th., 2003, Structure of linkage disequilibrium in plants, *Annual Review of Plant Biology*, 54:357~374
- Hudson R.R., 1985, The sampling distribution of linkage disequilibrium under an infinite allele model without selection, *Genetics*, 109: 611~631
- Hudson R.R., 2001, Two-locus sampling distributions and their application, *Genetics*, 159:1805~1817
- Jorde L.B., 2000, Linkage disequilibrium and the search for complex disease genes, *Genome Research*, 10:1435~1444
- Kruglyak L., 1999, Prospects for whole-genome linkage disequilibrium mapping of common disease genes, *Nature Genetics*, 22:139~144
- Lewontin R.C., and Kojana K., 1960, The evolutionary dynamics of complex polymorphisms, *Evolution* 14:458~472
- Long J.R., Zhao L.J., Liu P.Y., Lu Y., Dvornyk V., Shen H., Liu Y.J., Zhang Y.Y., Xiong D.H., Xiao P., and Deng H.W., 2004, Patterns of linkage disequilibrium and haplotype distribution in disease candidate genes, *BMC genetics*, 5:11
- Meuwissen T.H.E., Karlsen A., Lien S., Olsaker I., and Goddard M.E., 2002, Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping, *Genetics*, 161:373~379
- Nakajima T., Jorde L.B., Ishigami T., Umemura S., Emi M., Lalouel J.M., and Inoue I., 2002, Nucleotide diversity and haplotype structure of the human angiotensinogen gene in two populations, *The American Journal of Human Genetics*, 70:108~123
- Nothnagel M., Furst R., and Rohde K., 2002, Entropy as a measure for linkage disequilibrium over multilocus haplotype blocks, *Human Heredity*, 54:186~198
- Odani M., Narita A., Watanabe T., Yokoudi K., Sugimoto Y., Fujita T., Oguni T., Matsumoto M., and Sasaki Y., 2006, Genome-wide linkage disequilibrium in two Japanese beef cattle breeds, *Animal Genetics*, 37:139~144
- Pritchard J.K., and Przeworski M., 2001, Linkage disequilibrium in humans: models and data, *The American Journal of Human Genetics*, 69:1~14
- Przeworski M., and Wall J.D., 2001, Why is there so little intragenic linkage disequilibrium in humans, *Genetic Research*, 77:143~151
- Rao Y.S., Liang Y., Xia M.N., Shen X., Du Y.J., Luo C.G., Nie Q.H., Zeng H., and Zhang X.Q., 2008, Extent of linkage disequilibrium in wild and domestic chicken populations, *Hereditas*, 145:251~257
- Sabeti P.C., Reich D.E., Higgins J.M., Levine H.Z., Richter D.J., Schaffner S.F., Gabriel S.B., Platko J.V., Patterson N.J., McDonald G.J., Ackerman H.C., Campbell S.J., Altshuler D., Cooper R., Kwiatkowski D., Ward R., and Lander E.S., 2002, Detecting recent positive selection in the human genome from haplotype structure, *Nature*, 419:832~837
- Shifman S., Kuypers J., Kokoris M., Yakir B., and Darvasi A., 2003, Linkage disequilibrium patterns of the human genome across populations, *Human Molecular Genetics*, 12:771~776
- Slatkin M., 2008, Linkage disequilibrium--understanding the evolutionary past and mapping the medical future, *Nature Reviews Genetics*, 9:477~485
- Sutter N.B., Eberle M.A., Parker H.G., Pullar B.J., Kirkness E.F., Kruglyak L., and Ostrander E.A., 2004, Extensive and breed-specific linkage disequilibrium in *Canis familiaris*, *Genome Research*, 14:2388~2396
- Tiret L., Poirier O., Nicaud V., Barbaux S., Herrmann S.M., Perret C., Raoux S., Francomme C., Lebard G., Trégouet D., and Cambien F., 2002, Heterogeneity of linkage disequilibrium in human genes has implications for

- association studies of common diseases, *Human Molecular Genetics*, 11:419~429
- VanLiere J.M., and Rosenberg N.A., 2008, Mathematical properties of the r^2 measure of linkage disequilibrium, *Theoretical Population Biology*, 74:130~137
- Wall J.D., and Pritchard J.K., 2003, Haplotype blocks and linkage disequilibrium in the human genome, *Nature Review Genetics*, 4:587~597
- Wang J.P., Wang J.Q., Bu D.P., Liu N., Li F.D., and Luan S.Y., 2009, Review of physiological functions and metabolism mechanism of fat, *China Animal Husbandry and Veterinary Medicine*, 36:42~45 (王建平, 王加启, 卜登攀, 刘宁, 李发弟, 栾绍宇, 2009, 脂肪的生理功能及作用机制, *中国畜牧兽医*, 36:42~45)
- Wang R.H., Wang T.Y., and Li Y., 2007, Linkage disequilibrium in plant genomes, *Hereditas(Beijing)*, 29:1317~1323 (王荣焕, 王天宇, 黎裕, 2007, 植物基因组中的连锁不平衡, *遗传*, 29:1317~1323)

体外成功培养出小鼠有功能的精细胞

In vitro Production of Functional Sperm in Cultured Neonatal Mouse Testes

精子的生成是体内一系列细胞经过增殖和分化而来,这是一个复杂而漫长的过程。从精原细胞有丝分裂到精子形成要经过一个多月的时间。到目前为止整个过程还没有在除了几种鱼之外的其他物种体外复制出来。近期日本横滨市立大学研究院的科学家使用了转有 GFP(green fluorescent proteins)绿色荧光蛋白基因的小鼠作为研究对象,他们将新生小鼠中只含有生殖母细胞的睾丸组织消化为直径 1~3 mm 的碎片,放在半浸入培养基里的琼脂糖凝胶上面培养。根据绿色荧光的量来检测体外减数分裂的发生。起初他们使用的是小鼠出生后 7~10 d 的睾丸组织,培养后发现表达大量的绿色荧光。后来的实验发现,30~40 d 龄的小鼠睾丸组织碎片体外培养表达的绿色荧光蛋白量依然在最高水平,随着出生后时间的增长则开始逐渐减弱,但是可以持续到 70 d,说明精细胞的生成可以维持两个多月,而且 GFP 的表达并不影响精细胞的生成。流式细胞分析培养出的精细胞证明是单倍体的。中间他们更换了不同的培养基质,发现无血清的基质更有利于精子的生成。通过显微注射不同年龄新生小鼠体外产生的精细胞以及不同培养时间产生的精子,一共产生 12 只小鼠,在 3 周后断奶。PCR 检测发现有 4 只 GFP 阳性后代,将这些小鼠近交发现具有正常的生殖能力。此外他们将新生小鼠的睾丸组织在液氮中冻存了 4~25 d,经过解冻后进行了相同的验证,发现完全能够在体外生出精细胞。研究人员介绍这种器官培养方法经过改进后可以适用于很多哺乳动物,而且可以为临床应用和了解精子生成的机制提供一个平台。这项研究结果发表在 2011 年 3 月的《Nature》杂志上。

编者:黄甜(中国农业大学 农业生物技术国家重点实验室),本刊通讯员

本文引用格式:黄甜,2011,体外成功培养出小鼠有功能的精细胞,农业生物技术学报,19(2):286

信息来源:Nature doi:10.1038/nature09850